



Bulletin de méthodologie sociologique

Bulletin of sociological methodology

86 | 2005

April

Faire et défaire des groupes : L'information chiffrée sur les « populations difficiles à atteindre »

Marie-Ange Schiltz



Édition électronique

URL : <http://journals.openedition.org/bms/934>

ISSN : 2070-2779

Éditeur

Association internationale de méthodologie sociologique

Édition imprimée

Date de publication : 1 avril 2005

Pagination : 30-54

ISSN : 0759-1063

Référence électronique

Marie-Ange Schiltz, « Faire et défaire des groupes : L'information chiffrée sur les « populations difficiles à atteindre » », *Bulletin de méthodologie sociologique* [En ligne], 86 | 2005, mis en ligne le 28 mars 2008, consulté le 19 avril 2019. URL : <http://journals.openedition.org/bms/934>

Ce document a été généré automatiquement le 19 avril 2019.

© BMS

Faire et défaire des groupes : L'information chiffrée sur les « populations difficiles à atteindre »

Marie-Ange Schiltz

Introduction¹

- 1 Le souci de disposer de données de qualité sur les populations les plus démunies, en situation de précarité, est associé dès l'origine à la mise en place de l'enquête sociale (Bulmer *et al.*, 1991 ; Desrosières, 1993). Après guerre, les organismes ayant vocation à fournir de l'information statistique s'interrogent sur les « laissés pour compte » des recensements et des grandes enquêtes en population générale. Mais, ce n'est qu'au début des années 90 que l'interrogation sur les « marges » s'immisce en tant que sujet autonome dans l'activité statistique d'Etat.
- 2 L'intérêt récent des grands organismes statistiques ne doit pas faire oublier le souci opiniâtre d'une certaine recherche sociologique de rendre compte des modes de vie non conventionnels de groupes minoritaires, marginaux, en s'appuyant sur un arsenal de techniques qualitatives éprouvées. Parfois, en rupture avec cette tradition bien établie, l'enquête de « terrain » cherche, envers et contre tout, à rassembler des informations quantitatives sur des populations aux marges, étant bien entendu que, quelque soient les précautions de méthode mises en œuvre, aucune de ces tentatives bricolées ne peut prétendre à l'assise du raisonnement probabiliste.
- 3 L'irruption, dans le système statistique d'Etat, de l'interrogation statistique sur des groupes marginaux, mobiles, vivant dans les interstices de l'espace social, est corrélative au développement des politiques sociales en direction des populations les plus démunies et des gens vivant aux marges. Cette ingérence est soutenue par une opinion publique de plus en plus favorable au « traitement social » - secourable ou répressif - de la « question sociale » ou à la mise en place de « mesures compensatoires » pour contrebalancer certaines disparités ; elle est également confortée par un mouvement de revendication

des individus « différents » pour la reconnaissance du droit d'exister, du droit à l'indifférence ou du droit à la réparation. Dans ce contexte, une information de qualité s'impose désormais comme un préalable incontournable à la mise en œuvre éclairée des politiques sociales.

- 4 Le développement de la statistique d'Etat sur des sujets « sensibles » suscite actuellement une large réflexion sur l'acceptabilité des questionnements et la sincérité des réponses obtenues à propos de sujets délicats, à savoir des opinions politiques (Mayer, 2002), des comportements intimes, répréhensibles, illicites ou des situations douloureuses qui peuvent éventuellement déstabiliser et mettre en danger les personnes interrogées (Schiltz *et al.*, 2001). Ainsi, par quatre fois, au cours de ces dix dernières années, l'INED dans ses publications a ouvert ses colonnes à des réflexions sur la sincérité et la précision des réponses obtenues à propos de la sexualité, des enfants nés hors mariage, de l'illettrisme selon le contexte plus ou moins contraignant dans lequel se déroule l'interrogation : familial pour des adolescents consommant des drogues illicites, carcéral à propos de comportements répréhensibles.²
- 5 Alors que le débat sur la fiabilité des réponses obtenues à partir d'enquêtes quantitatives portant sur des thèmes « sensibles » est largement engagé (Rogers *et al.*, 1999), la réflexion sur la définition et les limites de ces populations souvent fragilisées, objet de ces interrogations délicates, est à peine effleurée. Dans le champ de la statistique, cette problématique est évoquée plus que traitée sous le vocable de population « aux contours flous, difficile d'accès », qui renvoie à un complémentaire tout aussi énigmatique, celui d'une « population cachée » qu'il s'agit d'éclairer. Dans le champ de la sociologie, ces questions sont abordées selon la théorie de la désignation, en termes de carrière de déviants, de stigmatisés et de discrédités ou de résistance au classement de la part d'individus qui maintiennent le secret sur leurs activités ou refusent l'aide dont ils pourraient bénéficier (Goffman, 1975 ; Becker, 1985 ; Castel, 1995a ; Rosanvallon, 1995). De ce fait, toute approche quantifiée des comportements non conventionnels se heurte au paradoxe du traitement d'une information - suspectée d'incomplétude, confiée dans l'intimité et l'anonymat de l'interview - qui doit être rendue publique³.
- 6 A ces divergences s'ajoute la tension entre l'exigence probabiliste qui suppose l'équiprobabilité d'inclusion des individus et la méfiance des sciences sociales à l'égard d'une approche qui établit l'équivalence entre une action et un classement social. A l'accusation d'empirisme émise à l'encontre de la standardisation des concepts et des techniques d'investigation répond celle de production de chiffres illustratifs sans aucune possibilité d'universalisation.
- 7 Enfin, la ligne de démarcation entre un discours fondé sur le classement qui se veut le plus objectif possible et une approche qui tente de prendre en compte la façon dont certaines désignations affectent les trajectoires et les modes de vie des individus, est irréductiblement au cœur du débat entre les statisticiens d'état et les sociologues de terrain. Alors que le statisticien a un problème de définition, le sociologue pense en terme de norme et de carrière.
- 8 Cette étude s'appuie sur une liste - présentée en annexe - d'enquêtes françaises référencées dans le texte selon leur sigle.

Définitions de la population « cible »

- 9 La réalisation d'une investigation auprès de populations difficiles à atteindre en raison de leur caractère minoritaire, de l'opprobre qui pèsent sur elles, de leur marginalité, de leurs conditions de vie, doit non seulement résoudre les problèmes d'accès à une minorité statistique mais aussi se confronter à la question des critères d'appartenance qui ne sont ni clairs, ni visibles. Ainsi, dans leur travail sur les « sans domicile fixe », groupe par essence mobile et difficile à identifier, M. Marpsat et J.M. Firdion (2000) proposent une première définition qui semble limpide : « sont sans domicile fixe les personnes qui ont 'perdu un foyer', celles qui, à une date donnée, passent la nuit dans l'espace public, dans un centre d'hébergement ». Or, leur travail de synthèse sur l'historique des connaissances accumulées sur les plus démunis parmi les plus démunis, montre que cette définition est incomplète ; en effet, à ces personnes peuvent s'ajouter, selon les enquêtes, celles qui passent la nuit chez des amis sans arrangement durable ou dorment épisodiquement à l'hôtel. La diversité des critères d'inclusion est révélatrice de l'hétérogénéité d'un groupe sans cohésion sociale, plus défini par le regard de la société que par les interactions entre ses membres (Damon, 2002). De même, une définition comportementaliste de l'homosexualité - avoir eu au moins un rapport sexuel avec une personne de même sexe au cours de sa vie - ne résout pas la délimitation des contours de la population « cible » en raison de la multiplicité des caractérisations possibles de ce qui est « homosexuel » : fantasmes, dispositions, attirance, pratiques sexuelles, affirmation de soi en tant qu'homosexuel, mode de vie ou identification à un groupe social et du flou qui entoure la notion de rapport sexuel (Pollak, Schiltz, 1991).
- 10 Ces deux exemples montrent ainsi qu'il existe une pluralité de définitions possibles des marges et qu'avant de délimiter, il faut définir. Or, tous les groupes minoritaires, non-conformes sont caractérisés par des « écarts de conduite » qui renvoient à une théorie de la « centralité », de la majorité, de l'individu « ordinaire »⁴, de la norme, de l'intégration, de la bonne conduite... Les qualifications de l'écart et de la « centralité » ne sont pas univoques ; les champs politique, administratif, juridique, statistique et sociologique mettent chacun en œuvre des logiques de classement selon des critères d'utilité et de compréhension qui leur sont propres.
- 11 Pour identifier la différence, une façon de faire consiste à faire référence à l'« essence » du groupe ; une autre consiste à s'intéresser aux relations entre des groupes et des individus inégaux.
- 12 Ainsi, I. Hacking (1986) rappelle que l'homosexualité peut être définie en référence à :
 - des comportements minoritaires (fantasmes, dispositions, attirance, pratiques sexuelles...) ;
 - une culture homosexuelle ; c'est-à-dire, les individus qui s'affirment comme « homosexuel », sont engagés dans un certain mode de vie gai ou s'identifient au groupe social ;
 - des gènes homosexuels ;
 - l'homosexuel en tant qu'espèce particulière de personne.
- 13 Ces classements ne sont pas exclusifs les uns des autres, une même marginalité peut être perçue sous des angles différents.
- 14 *Le point de vue statistique* définit les populations difficiles à atteindre :

- 15 en terme de probabilités plus fortes des individus d'échapper à l'activité statistique, que ce soit celle du recensement ou des grandes enquêtes. La pratique d'enquêtes post-censitaires⁵ permet d'évaluer ces biais d'échantillonnage (Coeffic, 1993) ;
à partir d'un dénombrement. La comparaison des effectifs permet alors de définir ce qui est minoritaire au regard de la majorité (les rouquins, par exemple) ;
en fonction de ce qui s'écarte par trop de la moyenne ou d'une autre valeur centrale (les obèses) ;
sur la base de réponses positives à un ou plusieurs comportements dits à « risque » définis préalablement (par rapport au risque du VIH : les homosexuels, les toxicomanes, les immigrés originaires des régions où le virus sévit de façon endémique) ;
sur la base d'un échantillon tiré à partir d'un enregistrement administratif ou d'une clientèle susceptible de bénéficier d'une action sociale (prisonniers, RMistes, séropositifs) ;
à partir d'un échantillon en boule de neige constitué à partir d'un individu désigné comme membre de la population cible.
- 16 *Le point de vue sociologique* s'ancre, quant à lui, dans la reconnaissance collective d'un problème social et de son traitement. La définition de la population cible s'établit à partir d'un attribut négatif.⁶ Les comportements non conformes sont construits en opposition à des « centralités » aussi diverses que le mode de vie du plus grand nombre, les conduites conventionnelles, le droit, ce qui est jugé sain, l'état de bonne santé... sachant qu'il n'y a pas de marge en soi, mais des modes de vie désignés comme problématiques par la collectivité⁷. Dans cette optique, il n'y a de position marginale que lorsqu'il y a norme. Tous les groupes sociaux instituent des normes, des attentes normatives, des stéréotypes qui sont intériorisés par une majorité qui les respectent et se charge éventuellement de les imposer à ceux qui ne les appliquent pas, à savoir, ceux qui sont aux marges du social, ceux qui apparaissent non conformes, ceux qui ne répondent pas aux attentes normatives, transgressent les normes, font des choses interdites... Selon Becker (1985), cette étrangeté, souvent disqualifiée, est toujours le résultat d'initiatives d'un autrui qui instaure et applique la norme. Pour devenir un objet de connaissance et d'intervention sociale, la différence doit se peupler : il faut alors découvrir, identifier, appréhender, prouver la culpabilité ou remarquer la différence et stigmatiser la non-conformité avant d'évaluer la taille du groupe.
- 17 En règle générale, l'approche sociologique s'intéresse moins à l'ampleur et aux causes intrinsèques de la marginalité qu'aux interactions sociales, au processus de marginalisation et aux effets des assignations négatives sur la perception de soi et les trajectoires de vie. En effet, selon la théorie de l'étiquetage, les carrières des individus ainsi catalogués relèvent autant du regard des autres que de leurs caractéristiques personnelles.
- 18 Il faut également trouver *les mots pour compter*, la valeur sémantique des mots étant d'autant plus difficile à clarifier que ces mots renvoient à des débats de société et des querelles sur les institutions : « L'accord sur le sens d'un mot est une forme instituée par la société, et est d'autant plus acquis que l'institution est consacrée » (Lévy, 1986).
- 19 Au début des années 1980, la réflexion menée à l'occasion de la révision de la nomenclature des professions (Desrosières, Thévenot, 1988) met en évidence la difficulté de classer pour compter. Cette difficulté s'impose alors à la communauté des statisticiens et la question de la délimitation des contours des populations, désormais, hante toute opération quantitative dont le but premier reste d'établir une mesure du

phénomène étudié (Hacking, 1989) : combien de sans domicile fixe (Marpsat, Firdion, 2000), combien d'homosexuels (Fay et al., 1989 ; Messiah, Mouret-Fourme, 1993), avant de savoir pourquoi et comment. La mobilisation politique et la transformation d'un problème social en « cause » ne sont pas indépendantes de l'ampleur estimée du phénomène (Jaspard et al., 2001). Une telle mesure détermine si le problème social est résoluble et quels sont les moyens à mettre en oeuvre, la nature de l'intervention politique différant du tout au tout selon le nombre de personnes concernées : un million, cent mille ou dix mille personnes ne se gouvernent pas de la même façon.

- 20 Le caractère minoritaire associé au refus de certains de se soumettre à une désignation négative avec, pour conséquence, l'absence de base de sondage ou de base de sondage complète, conduit à s'interroger sur la population de référence à partir de laquelle le chercheur construit l'échantillon : population éligible, population en demande, population susceptible d'une intervention sociale ou population effectivement prise en charge, population qui se range dans une classe. Le choix est toujours discutable. De récentes études ont examiné les groupes obtenus à partir de différents points de vue sur le handicap, la pauvreté ou les sans domiciles fixes. Les différents découpages donnent lieu à des configurations très diversifiées qui non seulement sont de grandeur différente, mais ne se recouvrent jamais complètement.
- 21 Même s'il est désormais acquis que ce « combien » n'est pas un absolu et que les contours de la population varient en fonction de la définition sémantique choisie (Ravaud *et al.*, 2002), la demande pour une évaluation des grandeurs du phénomène reste une question permanente posée au chercheur.

La lutte pour sortir ou redéfinir le classement

- 22 La délimitation des populations difficiles à joindre ne peut ignorer la lutte des individus classés à partir de comportements non conventionnels et dépréciés.⁸
- 23 Le projet d'une connaissance statistique se trouve confronté à des stratégies :
protectrices avec maintien du secret sur la base d'omission, de négation ou, au contraire, d'exagération d'une bonne présentation de soi ;
revendicatrices avec l'exigence d'abandon ou de redéfinition du classement, de demande de réparation.
- 24 La contestation de la désignation peut prendre plusieurs voies.
- 25 Ainsi, les représentants des minorités peuvent réclamer un dénombrement afin d'établir leur force sociale et politique ; certains peuvent même envisager sur le long terme la conquête de la majorité démographique et électorale.
- 26 Une plus grande fécondité des catholiques d'Irlande du Nord augure, dans un avenir proche, de la possibilité d'une parité démographique avec les protestants en Irlande du nord (Courbage, 1996).
- 27 Autre exemple, les homosexuels, forts des résultats de l'enquête Kinsey et du nombre mythique de 10% d'homosexuels que la société américaine des années 50 retient, prennent conscience de leur pouvoir. Ce sentiment d'appartenance libère leur parole et soutient la construction de leur identité collective et d'une communauté organisée (Lhomond, 1997).

- 28 Une autre stratégie de retournement du label consiste - en reprenant les termes de l'équipe de recherche du Cesdip (Robert *et al.*, 2003) - à « donner congé à l'homme moyen ».
- 29 Nonobstant la force du nombre, les groupes marginalisés peuvent alors :
réclamer le droit à l'indifférence ;
exiger l'abrogation de l'étiquetage ;
faire prévaloir leur droit de cité avec les mêmes droits et devoirs que le commun des mortels ;
mettre en valeur leur statut de victime afin de bénéficier d'aides publiques ou des avantages liés à une politique de discrimination positive.
- 30 Mais, ces stratégies revendicatrices ne sont pas accessibles aux populations les plus démunies qui ne partagent qu'une même « différence honteuse » ; trop atomisés, ils ne peuvent se regrouper en vue d'agir pour échapper ou retourner à leur profit la désignation négative.

Les principales stratégies statistiques pour délimiter la population « cible »

La défaillance de l'instrument scientifique

- 31 Les recensements et les grandes enquêtes représentatives se préoccupent des modes de vie de la population générale.⁹ Sur la base d'un projet d'exhaustivité ou de représentativité, le centre est considéré comme la totalité des répondants ou la partie représentant le tout. Dans cette optique, les marges, dites résiduelles, sont constituées de personnes ayant une probabilité plus forte d'échapper à l'activité statistique. Ces « reliquats » sont caractérisés par des enquêtes complémentaires au recensement ou des procédures de calage et de redressement des résultats statistiques.

L'enquête de contrôle du recensement de 1990 conclut à une probabilité d'omission plus forte en zone urbaine, dans les petits immeubles, les chambres d'hôtel, les constructions provisoires et les habitations de fortune. L'unité de base du recensement étant le logement, la population ayant une probabilité d'échapper au recensement est plus mobile (plusieurs résidences) dans une situation de transition (étudiants), professionnellement plus précaire et sans domicile fixe (Coeffic, 1993).

- 32 La marginalité ainsi définie est alors considérée comme un problème de collecte des données. Ce problème de collecte est d'autant plus important que la plupart des comportements non conventionnels renvoient à des caractéristiques sociales et à des styles de vie qui peuvent augmenter la probabilité des personnes de ne pas être contactées (Beck *et al.*, 2005). D'une certaine façon, la réflexion sur les instruments et les techniques de collecte fait partie de la construction des populations difficiles à atteindre : l'enquête SDF (réalisée auprès des personnes sans domicile fixe) ne naît-elle pas de la prise de conscience des limites des enquêtes « ménages » menées par l'Insee.

Le travail de classement

- 33 Les enquêtes surdimensionnées de filtrage en direction de minorités stigmatisées opèrent à partir d'une base de sondage constituée en population générale. La délimitation des groupes « cibles » s'effectue a posteriori sur la base de comportements prédéfinis

(enquête Enveff) ou à partir d'une (enquête ACSF) ou plusieurs (enquête HID) questions filtres. Le recrutement spécifique est usuellement complété par celui d'un échantillon témoin sélectionné selon une procédure aléatoire.

- 34 En général, dans ce genre d'enquête, la ligne de conduite est de rester discret sur le sujet sensible de l'enquête et de ne le dévoiler que progressivement.

Dans la "lettre annonce" à entête de l'Inserm, l'enquête ACSF est présentée comme « une grande étude sur la santé auprès de 20 000 personnes ». Au moment du premier contact téléphonique, il est annoncé que cette « grande étude sur la santé » est destinée « à mieux définir la prévention » ; c'est seulement lorsque l'individu est sélectionné que le sujet complet est dévoilé, l'enquêteur annonce alors que « cette étude porte sur la santé et la prévention du sida, notamment sur vos opinions, sur ce que vous connaissez de cette maladie, votre vie sexuelle et sur ce que vous pensez des mesures de prévention » (ACSF, 1993 : 1242)

Dans l'enquête Enveff, les auteurs précisent que « Le sujet des violences envers les femmes demande à être introduit de façon particulière et délicate, afin de ne pas induire de réaction de refus de la part de ménages ou de femmes qui ne se sentiraient pas du tout concernés ou, au contraire "trop" concernés et qui pourraient craindre les conséquences d'une telle enquête. L'envoi d'une "lettre annonce" favorise généralement un meilleur taux de réponse. Si cette lettre met l'accent sur l'importance de l'étude, elle doit être motivante et rassurante sans trop dévoiler le sujet de l'enquête. Après de nombreuses réflexions, le choix de l'intitulé s'est finalement porté sur « les conditions de vie, la santé et la sécurité en France » ; le sujet de la sécurité ayant été jugé plus porteur auprès de l'opinion publique que celui de la violence qui pouvait éveiller la suspicion d'un proche violent. Pour cette même raison, la spécificité de la population étudiée n'a jamais été précisée, des femmes uniquement. »

- 35 De la même façon, la rédaction d'une question filtre doit rester suffisamment vague afin d'augmenter les chances de recruter des personnes qui ne veulent pas parler.

Ainsi, dans l'enquête ACSF (Analyse des comportements sexuels en France)¹⁰, le choix d'un questionnaire court (socio-démographique) ou long (socio-démographique et comportements sexuels) est conditionné par une réponse positive globale à une question rédigée de telle façon que le répondant n'a pas à énoncer une identité sexuelle ou une conduite répréhensible qui peut être difficile à déclarer.

Je vais vous lire une liste de situations et vous me direz juste si une de ces situations vous concerne, sans me préciser laquelle, je dis bien sans me préciser laquelle :

Vous êtes né (e) le 4, le 17 ou le 20 d'un mois de l'année

Vous avez eu des rapports sexuels avec au moins deux personnes différentes dans les 12 derniers mois

Au cours des 5 dernières années vous avez eu des rapports sexuels au moins une fois avec une personne du même sexe que vous

Au cours des 5 dernières années, vous avez payé au moins une fois pour avoir des rapports sexuels

Au cours des 12 derniers mois vous avez utilisé au moins une fois une drogue douce (hasch, marijuana,...) ou dure (cocaïne, héroïne,...)

Vous êtes hémophile

Ne veut pas répondre

Toute personne ayant répondu positivement à cette question filtre est soumise au questionnaire détaillé. 4820 personnes ont ainsi été interrogées. Au final, l'échantillon est composé à parts presque égales d'individus « témoins » (être né...) et d'individus relevant d'un critère de risque face au VIH.

- 36 Dans la même logique, il s'agit d'interroger sur des comportements pré-définis que l'on évite de nommer.

Dans l'enquête *Enveff*, l'interrogation sur la violence s'effectue sans jamais la "nommer" : « nous avons évoqué uniquement des gestes, actes, faits, paroles, susceptibles de porter atteinte à l'intégrité physique et morale de l'autre sans les qualifier de violents » ;

Dans l'enquête *Acsf*, les homosexuels ont été sélectionnés « sans être questionnés directement sur une identité sexuelle dont on sait à quel point elle peut être difficile à dire, voire indicible ». (ACSF, 1993 : 1355)

- 37 Les groupes statistiques « ciblés » sont construits a posteriori sur la base de réponses positives aux unités de conduite préalablement sélectionnées ou d'un risque plus fort attaché à certains groupes d'individus.¹¹ Les caractéristiques des groupes ainsi repérés sont interprétées au regard de celles de la population inclus dans l'étude n'ayant pas les attributs requis, souvent qualifiée de « population ordinaire ».

- 38 D'autres procédures de filtrage plus complexes prennent en compte différentes dimensions des classements. Ainsi, l'opération de filtrage de l'enquête HID (handicaps-incapacités-dépendance) menée auprès des ménages par l'Insee opère tout à la fois à partir de comportements, de besoins d'assistance, d'expertise, d'une reconnaissance administrative et d'une auto-désignation des personnes (Ravaud et al., 2002).

Le premier volet de cette interrogation s'appuie sur une enquête complémentaire au recensement de 1999 sur le thème « vie quotidienne et santé ». Ce fichier servant, par la suite, de base de sondage pour l'enquête HID proprement dite. Un filtrage multidimensionnel correspondant à cinq approches possibles du handicap autorise différentes constructions de la population handicapée qui s'ordonnent selon leur ordre de prévalence de la façon suivante :¹²

les incapacités fonctionnelles appréhendées par 7 questions factuelles sans référence explicite aux dénominations de handicap, d'invalidité ou de dépendance : 21,1% ;

Une question générale sur la limitation dans le genre ou la quantité d'activités en prenant en compte l'aspect durable de la difficulté (plus de 6 mois) : 11.7% ;

L'auto-déclaration du handicap : 9.5% ;

Le recours à une aide : humaine, prothèse, appareil, technique, aménagement du logement : 7.7% ;

Demande de reconnaissance sociale officielle du handicap : 6.2% (dont une demande sur six refusée ou en attente de réponse) ;

- 39 Notons que chacun de ces groupes ainsi construits sont de tailles extrêmement différentes mais qu'en plus ils ne se recouvrent pas complètement.

La désignation sociale

- 40 Les statisticiens ont développé des procédures de sélection aléatoire en prenant comme base de sondage une population touchée par une réponse sociale spécifique. Dans ce cas, la construction du groupe statistique s'appuie sur les critères de l'intervention sociale ; il est alors possible d'imaginer des procédures de constitution d'échantillons représentatifs construits à partir de la gestion ou de l'offre de service de la réponse sociale.

- 41 Certaines approches utilisent comme base de sondage les fichiers administratifs (*Enquête RMI*) ou les listes des administrés (*Enquête Histoire familiale, volet hommes incarcérés*). Quand ces fichiers n'existent pas, une alternative de sondage est envisageable sur la base des prestations individuelles. Dans ce cas, pour constituer un échantillon représentatif, il est nécessaire d'établir une base de sondage exhaustive qui se présente sous la forme de la liste complète des unités prestataires fréquentées par la population visée, liste à partir de laquelle un premier tirage au sort est effectué.

Pour établir leur base de sondage, l'enquête *Histoire familiale, volet hommes incarcérés* menée par l'Insee en 1999, dresse d'abord la liste exhaustive des établissements pénitentiaires de France métropolitaine (Cassan, 2002)¹³.

L'enquête *SDF-Ined* dresse, quant à elle, la liste exhaustive des services d'hébergement et de distribution de nourriture à Paris ;¹⁴ le tirage des centres se faisant proportionnellement à leur capacité d'accueil.

L'enquête *Vespa* se propose de réaliser une enquête auprès d'un échantillon représentatif de la population ayant connaissance de son infection par le VIH, sachant que la grande majorité des personnes atteintes consultent en milieu hospitalier¹⁵. De ce fait, l'enquête se concentre sur l'activité hospitalière. La liste exhaustive des services prenant en charge la pathologie est d'abord établie ; elle sert de base de sondage au tirage des services stratifiés en fonction de la taille de leur « file active » de patients et de la région.

- 42 Puis, dans chaque unité prestataire ainsi sélectionnée, on procède à un tirage aléatoire des individus ou des prestations.

L'enquête *Histoire familiale volet « détenus »* effectue un tirage aléatoire parmi les détenus présents dans l'établissement sélectionné.¹⁶

L'enquête *SDF-Ined* réalise un tirage aléatoire de six prestations par service sélectionné.

L'enquête *Vespa* procède, quant à elle, à un tirage aléatoire des consultations auprès de chaque médecin consultant dans le service sélectionné.

- 43 Lorsque l'échantillonnage s'effectue à partir des prestations, une opération supplémentaire est nécessaire afin de passer de la prestation à l'individu bénéficiaire. Ce glissement implique un questionnement rétrospectif sur la fréquence d'utilisation des services par l'individu interrogé afin de pouvoir établir des pondérations qui tiennent compte de la probabilité de sa présence dans l'unité sélectionnée au moment de l'enquête

¹⁷.

L'identité revendiquée

- 44 Plutôt que de décider d'un comportement à partir duquel les individus sont déclarés « étrangers » au regard des comportements conventionnels, le chercheur peut opter pour la constitution d'un fichier de données quantitatives sur la base du volontariat. Dans ce cas, les critères de l'inclusion lui échappent pour ne plus dépendre que de la décision des individus qui :

se classent volontairement en réponse à une question explicite -
auto-positionnement sur une échelle hétérosexualité-homosexualité dans l'enquête Kinsey en 1946 : de 0 exclusivement hétérosexuel à 6 exclusivement homosexuel ;
auto-classement comme handicapé dans l'enquête HID en 1999.

décident de s'inclure dans une enquête dont le titre indique, cette fois-ci, clairement son objet, à savoir l'étude de modes de vie non conformes -

homosexualité : *Rapport sur l'homosexualité de l'homme* réalisée auprès du public d'une association et de sa revue (Bon, d'Arc, 1974); *Rapport Gai* réalisé dans le milieu militant et associatif (Cavailhès et al., 1984); *Enquêtes Presse Gay* répétées pendant plus d'une décennie auprès de lecteurs de la presse homosexuelle (Pollak, Schiltz, 1991 ; Schiltz, 1997) ;
les autodidactes : enquête à travers la revue *Science et vie* (Boltanski, Maldidier, 1977) ;

- 45 En choisissant cette forme de recrutement, le chercheur estime qu'il peut prendre au sérieux ces individus qui, par leur auto-désignation ou leur participation volontaire, se réclament d'une certaine appartenance à une communauté.¹⁸ D'un certain point de vue, il est, en effet, pertinent de s'intéresser à celui qui se perçoit comme stigmatisé, déviant et qui, maintenant sur une longue période une forme déterminée de déviance ou de stigmatisme, organise sur cette base son identité et son mode de vie, plutôt qu'à celui qui commet un acte exceptionnel.
- 46 Reste alors en suspend, le problème du recrutement des « volontaires ». Comment faire savoir l'intention d'une enquête sociale, susciter le témoignage, inciter à rompre le silence ? Dans le cas d'une communauté qui se donne les moyens de s'affirmer et de s'afficher, le chercheur peut envisager de passer l'information par les formes institutionnalisées du groupe - associations et éventuellement presse communautaire - qui marquent sa visibilité et en pérennisent l'existence.

L'identité relationnelle

- 47 Les enquêtes à partir d'échantillonnage en boule de neige, sur quota ou non, renvoient à la notion d'identité relationnelle, chaque individu contacté donne accès à d'autres individus qu'il sait avoir la caractéristique recherchée. Dans ce cas, le recrutement de populations marginales s'apparente à la démarche usuelle des ethnologues qui se servent d'"informateurs privilégiés" et procèdent par "boule de neige".
- 48 Contrairement aux enquêtes auprès de volontaires, ce recrutement de proche en proche s'appuie sur l'existence de réseaux constitués à partir de pratiques et de connaissances qui peuvent être complètement informels. De ce fait, cette technique d'inclusion présente l'avantage de ne pas nécessiter de forme organisée du groupe ; permettant d'aborder des populations socialement peu organisées, cette approche est très utilisée pour l'exploration des réseaux les plus marginaux - ceux qui se constituent à partir de conduites délictueuses comme les consommations de substances psycho-actives - ou l'étude des populations « cachées » que sont les laissés-pour-compte des autres approches que nous avons examinées.

Limites de chaque type d'approche

- 49 L'absence de base de sondage, tout au moins complète, génère des difficultés à définir :
- la population de référence : population générale, population vulnérable, population éligible,¹⁹ population en demande, population susceptible d'une intervention sociale ou population effectivement prise en charge.
 - et son envers, la population « cachée » : population qui refuse le classement, rejetée par l'intervention sociale, qui adopte des stratégies alternatives, qui ignore ses droits, sa contamination...
- 50 En plus des biais classiques de recrutement, les enquêtes, qui sélectionnent la population cible sur la base d'un ou plusieurs filtres, sont confrontées à des problèmes liés à la taille des échantillons, au respect de l'équiprobabilité d'inclusion des individus et à l'exclusion à tort d'individus .

- 51 Toute enquête de filtrage nécessite un échantillon bien plus important que la population que l'on souhaite décrire, ce qui implique des coûts et des temps d'observation très importants.

Dans l'enquête ACSF, pour le volet « ménages », réalisée en 1991-92 auprès de 20 055 personnes, le rapport entre les individus ayant accepté de répondre et ceux interrogés de façon détaillée est d'environ un pour dix, si l'on s'en tient aux individus relevant d'un des critères de risque quant à la transmission du sida (2187), et de un à cinq si l'on inclut l'échantillon « témoin » composé de 2332 personnes.

Dans l'enquête HID-ménages 1999, ce rapport est de l'ordre de 1 à 20 : 22 000 personnes sélectionnées pour le questionnaire HID proprement dit, par rapport aux 417 500 personnes sollicitées pour répondre à l'enquête VQS (vie quotidienne et santé).

- 52 Il s'agit, en effet, de s'assurer des chances de contacter un nombre d'individus suffisamment important pour justifier d'une analyse statistique valide du groupe. Ce qui n'est pas toujours le cas.

Au total, dans l'enquête ACSF, 210 hommes ont déclaré au moins un rapport homosexuel au cours de leur vie et 103 au cours de l'année. En revanche, un effectif trop restreint de femmes homosexuelles n'a pas permis une analyse statistiquement valide du groupe.

Pour pallier cette éventualité, l'inclusion des individus dans l'échantillon HID s'effectue à partir de probabilités croissantes avec la sévérité présumée du handicap dans le but d'avoir un effectif conséquent de personnes lourdement handicapées.

- 53 L'hypothèse d'équiprobabilité d'inclusion des individus est parfois difficile à assurer. Le mode de vie des membres du groupe « cible » doit être comparable à celui de la majeure partie de la population et compatible avec les conditions de collecte des données. Or, cette hypothèse est parfois difficile à soutenir. En effet, dans chaque enquête, l'analyse des biais d'échantillonnage montre que certaines catégories de personnes sont plus difficiles à joindre, sachant que les modes de vie spécifiques de ces sous-groupes ne sont pas toujours sans lien avec le thème de l'enquête.

Dans l'enquête ACSF, des biais concernant la participation des homos et bisexuels sont possibles sans que l'on puisse savoir dans quel sens ils peuvent jouer et se compenser. La sur-représentation des plus diplômés peut jouer en leur faveur ; au contraire, la recherche de l'anonymat peut les conduire à s'inscrire sur la liste rouge qui n'est pas prise en compte par la base de sondage. Des modes de vie particuliers (célibat, fréquentation de la scène gaie, sorties, nombre plus élevé de partenaires) ne sont pas sans lien avec une difficulté plus grande de contacter ces personnes par téléphone.

Dans l'enquête Enveff, on peut se demander de quelle façon se répercute la difficulté de joindre par téléphone les femmes jeunes et célibataires sachant que ce groupe a un degré plus fort d'exposition au risque de violence. Et parmi les femmes qui ont échappé à l'enquête se trouvent des groupes probablement fortement touchés par les phénomènes de violence comme celles hébergées en foyer, en centre d'accueil d'urgence ou sans domicile.

- 54 Le groupe statistique constitué, à partir des réponses positives à une ou plusieurs questions « filtres », exclut les « faux négatifs », à savoir ceux qui n'ont pas répondu positivement aux questions filtres - par erreur, omission, refus de le dire, refus de se reconnaître dans la catégorie.²⁰

La comparaison des résultats des enquêtes Cездip-IdF ((Pottier et al, 2002) et Enveff (Jaspard et al., 2001), toutes deux réalisées en population générale en vue de produire des données valides sur divers types de violences, fournit des évaluations systématiquement beaucoup plus élevées pour l'Enveff. La disparité des ordres de

grandeurs, souvent dans un rapport de 1 à 4, se répète pour tous les indicateurs susceptibles de comparaison. Pour n'en citer qu'un, la fréquence des agressions sexuelles à l'encontre de femmes âgées de 20 à 59 ans vivant en Ile de France au cours de la dernière année est, selon l'enquête Cездip-IdF, de 0.28%, et de 1.2% dans l'enquête Enveff. Cette différence peut s'expliquer en partie par la différence de stratégies de filtrage. Tandis que le questionnaire Cездip-IdF s'appuie sur une seule question filtre qui se réfère explicitement à la notion d'agression pour évaluer l'ampleur du phénomène, la construction du questionnaire Enveff s'organise autour de la mise en situation des répondantes dans les différentes sphères de leur vie (publique, professionnelle, familiale, conjugale,...). Pour chacun de ces univers, une liste de faits se répète sans jamais que les mots d'agression ou de violence ne soient prononcés. Dans cet exemple, l'absence de filtrage versus un filtrage, des formulations différentes, la répétition des questions en référence à des sphères spécifiques d'activité destinée à favoriser le travail de mémoire et des niveaux d'investigation plus ou moins détaillés délimitent des populations de grandeur très différentes.

- 55 On sait désormais que le rendement d'un filtre dépend :
- de la formulation de la question selon qu'elle renvoie à des comportements, un classement administratif, à un auto-classement des personnes ;
 - de l'amplitude du champ exploré (niveau de détail du questionnement).
- 56 La constitution d'une connaissance à partir de fichiers administratifs et de populations « captives », ou sur la base d'un tirage de prestations individuelles, subodore une reconnaissance implicite de l'adéquation du système de prise en charge aux besoins des personnes concernées ou à l'accès aux services - de tous et des seuls - ayants droits. En cas de doute, cette approche pose un problème de caractérisation et de délimitation des populations cachées : ceux qui ont échappé à la désignation des services, les ayants droits ne revendiquant pas leur droit, ceux qui refusent la désignation attachée à la prise en charge, ceux qui ont fait une demande mais qui ont été déboutés.
- 57 Généralement, une investigation est conduite auprès des populations « cachées » afin de déterminer les caractéristiques de cette population « fantôme » par rapport à celles de la population enquêtée. Ce complément d'enquête permet de qualifier :
- les biais liés aux procédures de sélection des institutions ;
 - les caractéristiques des individus pris en charge, opposées à celles de ceux qui résistent ou échappent au traitement social ou médical.
- 58 Les enquêtes auprès de volontaires ne sont possibles que s'il existe une communauté organisée et le sentiment d'un destin commun qui s'exprime au travers d'associations ou d'une presse spécialisée dans lesquelles se reconnaissent les personnes concernées. L. Boltanski et P. Maldidier (1977) avaient remarqué que l'ensemble statistique de ceux qui se classent ou s'incluent volontairement selon un attribut négatif n'est pas le modèle réduit d'une assemblée plus vaste mais un concentré souvent déformé de la forme la plus durcie du noyau d'intérêts qui se regroupe autour de la marginalité, du stigmat exploré. Plus la définition est restreinte, plus les canaux de diffusion de l'enquête sont spécifiques, plus l'échantillon spontané présente des traits distinctifs par rapport à la population de référence.

La comparaison entre le recrutement du Rapport Gai qui recrute, en 1984, selon la méthode en boule de neige en utilisant les lieux de rencontre et les associations militantes, les enquêtes Presse Gay qui, de 1985 à 2000, s'appuient sur un questionnaire diffusé par voie de presse et l'enquête ACSF qui procède de façon aléatoire montre que :

- L'échantillon du Rapport gai se cristallise autour de la forme la plus durcie du

groupe avec une très forte affirmation de soi en tant qu'homosexuel, un écrasement de la pyramide des âges autour du jeune adulte, une concentration de la résidence dans les grandes villes et la région parisienne.

- Par rapport à cette dernière, l'enquête par voie de presse joint mieux les homosexuels provinciaux, plus jeunes et plus âgés.

- Par rapport à l'échantillon aléatoire, les volontaires de l'enquête Presse Gay sont essentiellement l'image restreinte des hommes qui déclarent des pratiques exclusivement homosexuelles dans l'année, une minorité parmi les hommes qui ont déclaré au moins un rapport homosexuel au cours de leur vie dans l'enquête ACSF.

- 59 Etant donné que l'entrée dans un groupe déviant organisé constitue souvent la dernière étape d'une carrière marginale, la diffusion d'une enquête par le biais des organisations communautaires a toutes les chances de donner l'image la plus durcie du groupe.
- 60 Cette forme de biais se retrouve de façon accentuée dans l'échantillonnage en boule de neige. En dépit d'une tradition de consignes strictes dispensées aux enquêteurs qui donnent une impression de contrôle de l'inclusion, ce mode de recrutement n'est pas toujours d'une rigueur impeccable. Sachant que, dans les sub-cultures, les groupes les moins marginalisés sont aussi les moins visibles, la procédure qui cherche à atteindre les membres d'un groupe par l'intermédiaire du réseau relationnel a toutes chances de ne toucher que la partie la plus visible et la plus affirmée de la population cible.
- 61 De fait, le recrutement en « boule de neige » met en œuvre une succession de filtres implicites ; quel que soit leur degré de familiarité du terrain, une partie des relations de réseaux - dans lesquels se retrouvent des populations marginales qui ont l'habitude d'évoluer dans le non-dit ou le secret - échappent à l'enquêteur même le plus indigène. Aussi, généralement, l'enquêteur est obligé de s'appuyer sur des informateurs - souvent les personnes les plus visibles du réseau - qui vont l'orienter vers d'autres personnes qu'ils savent avoir la caractéristique recherchée ; pour ce faire, il est nécessaire que l'individu recherché se soit suffisamment manifesté pour avoir été repéré par l'informateur qui, parmi d'autres possibles, le désigne à l'enquêteur. Cette succession de contrôles informels a toutes chances de rendre inévitables de forts biais de sélection au profit des individus les mieux incérés dans le réseau représentant le noyau le plus actif et le plus assidu. D'autant plus que les biais de sélection se trouvent renforcés par l'interaction avec l'enquêteur ; l'absence d'anonymat dans le face à face peut écarter de l'interrogation les personnes qui nient l'évidence, ne s'affichent pas ou encore celles qui ordinairement refusent de reconnaître devant un « étranger » leur comportement stigmatisé.
- 62 Ce type de biais est attesté par une étude comparative sur le recrutement de populations homosexuelles qui montre que le recrutement de proche en proche donne lieu à un échantillon encore plus distinctif du groupe « cible » qu'un échantillon constitué sur une base volontaire. Du fait de l'opprobre qui subsiste sur les conduites homosexuelles, la morphologie de la population homosexuelle qui entre dans une enquête ne se résume pas aux seules variables socio-démographiques (âge, diplôme, CSP, lieu de résidence), le degrés d'acceptation par soi-même et par autrui d'une disposition homosexuelle ou des taux de séropositivité donnent une autre image de la composition des échantillons et permettent d'établir l'évaluation du degré d'ouverture du recrutement.

En comparaison avec d'autres procédures de recrutement en France et à l'étranger, les enquêtes lecteurs attestent un degré d'acceptation perçue de l'homosexualité plus faible que celles aux répondants recrutés par d'autres voies. [...] ; un indicateur supplémentaire renforce notre conviction du meilleur recrutement par voie de presse dans la population homosexuelle qui ne s'affirme pas et qui reste peu

"visible" (à condition bien évidemment de l'existence et de la vente suffisante de la presse gaie dans un pays donné) : les enquêtes lecteurs, en France et ailleurs, témoignent toutes d'une séroprévalence de l'infection par le VIH légèrement plus faible que les enquêtes quotas et "boule de neige". Or, toutes les analyses épidémiologiques constatent que le virus a commencé à circuler dans les classes moyennes supérieures parmi les homosexuels s'identifiant le plus avec leurs pratiques qu'ils vivent ouvertement et que d'autres sont restés à l'abri du risque. Par conséquent, plus la séroprévalence est élevée dans un échantillon, moins il est représentatif de la population homosexuelle dans son ensemble (Pollak, Schiltz, 1991).

- 63 Par ailleurs, il faut souligner que ce mode de contact privilégie les lieux d'échanges (lieux de rencontres pour les homosexuels ou sites de rue ou institutionnels pour les toxicomanes) qui seront pour des raisons d'accessibilité la plus part du temps les plus visibles, les plus connus. A moins d'avoir recours à des techniques de saturation du réseau exploré, tout comme le recrutement de « volontaires », cette démarche ne procure aucune représentativité.

Les nouveaux développements

- 64 Depuis le début des années 1990, la volonté de traitement de la « question sociale » a insufflé des moyens techniques et financiers considérables qui ont permis des innovations méthodologiques en vue d'améliorer la qualité de l'information à propos des populations difficiles à atteindre.
- 65 Ces développements s'inscrivent principalement dans le cadre de la démarche inférentielle. Le surdimensionnement d'enquêtes représentatives, le tirage d'échantillon à partir de fichiers administratifs ou auprès de populations captives, les techniques de croisement de plusieurs sources de données permettent de réintroduire les règles de l'observation probabiliste.
- 66 Les enquêtes selon la méthodologie de saisie avec recoupement (« capture et recapture ») font partie de ces récents développements probabilistes qui, à partir d'appareillement de diverses sources d'information quantifiées, autorisent une estimation indirecte des prévalences avec calcul de l'intervalle de confiance²¹. Cependant, la transposition de cette technique à l'étude de la prévalence de comportements marginaux et stigmatisés comme la toxicomanie ou la prostitution n'est pas simple (Leyland et al., 1993). Dans ce cas, la constitution de l'échantillon « est affectée par des problèmes de définition et de recoupement qui n'existent pas dans les études de populations animales » (Richardson, 1996). En effet, la « capture-recapture » s'effectue non pas sur la base d'une procédure de capture aléatoire mais souvent sur la base de sources d'informations administratives en constante évolution et pas toujours indépendantes ; de plus, les probabilités des individus de figurer sur une liste varient selon leurs caractéristiques individuelles, les contextes de vie et les conditions d'approches des populations visées.

Ainsi l'étude multicentrique des estimation locales de la prévalence de l'usage d'opiacés en cocaïne en France (Chevalier, 2001) - dont les deux premières estimations de prévalence datent de 1994 et 1995 - mobilise des structures de prise en charge des toxicomanes aussi diverses que les lieux d'aide à la vie quotidienne des toxicomanes actifs en tenant compte de leur dépendance (accueil bas seuil d'exigence), les centres de soins, d'hébergement, les points d'écoute, les services de soins hospitaliers spécialisés ainsi que les sources répressives et judiciaires.

- 67 Dans ce cas, l'équiprobabilité d'inclusion des individus, l'indépendance et la fermeture des listes qui fondent la qualité de l'estimation ne sont pas assurées.
- 68 Les développements techniques de la collecte et de la mise en rapport de diverses sources de données permettent à l'heure actuelle d'envisager de prendre en considération les différents rapport au monde social des groupes minoritaires ou marginaux qui, selon R. Chartier (1989), se décomposent en opération de classement, façons de se manifester et formes institutionnelles –internes ou externes au groupe – qui marquent de façon visible l'assignation sociale. Ce souci de s'informer selon les différentes dimensions se matérialise dans les nouveaux développements de l'approche quantitative des populations en marge.
- 69 En ce qui concerne les enquêtes de filtrage surdimensionnées, il s'agit désormais de choisir de ne pas choisir soit en refusant de filtrer soit en prenant en compte la multiplicité des désignations possibles qui déterminent des groupes à géométrie variable afin de minimiser les classements négatifs définitifs.

Le questionnaire de l'enquête Enveff construit dans l'optique de favoriser la mémorisation et de faire émerger des actes de violences, soumet, sans aucun filtrage, une liste de faits de violence plus ou moins identiques qui se répètent dans les différentes sphères de la vie des femmes.

L'enquête VQS-HID prend résolument en compte le caractère flou des contours de la population en abandonnant « l'idée du tri impératif entre positifs et négatifs pour la construction de l'échantillon de deuxième phase ». (Morniche, 1998). Si, au stade de la reconnaissance administrative - « avoir ou pas un handicap », « être ou ne pas être handicapé » - sont des états qui partagent sans ambiguïté la population entre bénéficiaires d'un classement et d'une aide, et les autres, dans la réalité les choses sont plus équivoques. L'enquête HID multiplie les points de vue : à côté d'un ensemble de questions factuelles sur les incapacités fonctionnelles et l'aide nécessaire, et une question générale sur la limitation des activités (grilles internationalement en usage pour l'évaluation du handicap), il est demandé au répondant d'évaluer le désavantage qu'il ressent et ses éventuelles démarches en vue d'une reconnaissance officielle du handicap.

- 70 Le substrat multidimensionnel du filtrage détermine un groupe statistique à géométrie variable et autorise différentes constructions de la population cible. Il permet de considérer les recouvrements ou, au contraire, les dissonances entre les classements des chercheurs, des experts, des institutions et, enfin, de la personne concernée, et d'établir ainsi des bases solides pour une discussion sur la construction sociale du groupe. Il semblerait, par ailleurs, le taux de non recouvrement soit un indicateur intéressant de structuration du groupe au sens où ce taux est d'autant plus important que le groupe est controversé et faiblement constitué.
- 71 Par ailleurs, de plus en plus, grâce aux possibilités de l'informatique et des logiciels qui facilitent la gestion des bases de données, l'étude de populations ou de comportements marginaux est éclairée par le regard croisé de plusieurs techniques d'observation. Les protocoles d'enquête peuvent désormais multiplier les angles d'approches sur un même thème ou une même population, couplant des informations issues d'entretiens, de questionnaires fermés, de fichiers pré-établis sur les personnes et autres sources d'information, ou appliquant les mêmes outils de collecte de données à divers segments de la population afin d'éclairer les informations recueillies sur un groupe spécifique.

L'enquête Histoire familiale multiplie les volets d'une même enquête auprès de populations différentes ; concomitamment au bulletin complémentaire distribué à l'occasion du recensement auprès des « ménages ordinaires », en 1999, pour la

première fois, l'Insee, promoteur de cette enquête, a interrogé un échantillon d'hommes détenus sur le même thème et à partir du même outil d'interrogation, à quelques adaptations à l'univers carcéral près. Grâce à la similarité des questions, l'histoire familiale des hommes détenus est ainsi éclairée par celle des « citoyens libres ».

- 72 L'interrogation sur les caractéristiques de la population cachée s'impose désormais parmi les règles de bonne conduite d'une collecte de données.

Pour l'enquête VESPA, l'idéal aurait été d'enquêter auprès un échantillon aléatoire représentatif de l'ensemble de la population ayant connaissance de son statut VIH. Or, on sait que les séropositifs au fait de leur infection peu ou pas suivis à l'hôpital sont difficilement accessibles. Pour pallier cette carence, une série d'investigations auprès de cette population est entreprise afin d'éclairer l'information recueillie auprès des seules personnes atteintes suivies à l'hôpital. Il s'agit de dessiner les contours de la population « cachée » afin d'identifier les éventuelles possibilités de l'inclure dans une enquête quantitative. Les informations ainsi recueillies tendent à confirmer qu'il s'agit d'une population minoritaire et hétérogène qui ne peut faire l'objet d'une interrogation standardisée.

- 73 De nouveaux protocoles de collecte des données cherchent à minimiser l'ampleur de cette population cachée. Pour n'en citer que quelques uns, on peut évoquer les techniques d'échantillon raisonné, de saturation de l'espace investigué par capture et recapture ou de recrutement en boule de neige sous quotas.

- 74 Enfin, reprenant à leur compte la notion de trajectoires de vie, de carrière, ces nouvelles approches se préoccupent de plus en plus de distinguer l'acte exceptionnel, la situation accidentelle d'une vie façonnée par le comportement déviant, le handicap ou l'exclusion.

- 75 Les derniers développements de la statistique d'Etat montrent que l'approche statistique comme instrument de connaissance des individus les plus démunis, les plus fragiles n'est pas incompatible avec le souci de raisonner au plus juste dans l'incertain ; c'est-à-dire, dès que l'on en a les moyens de s'appuyer sur l'art du calcul probabiliste. Cependant l'ampleur de ces développements techniques ne doit pas masquer les interrogations qui restent en suspend.

- 76 Suivant Ian Hacking (1986), il s'agit de s'interroger sur ce qu'est une « population » et le problème de sa délimitation, tout en prenant en compte la préoccupation de Robert Castel (1995b) qui s'inquiète de la distorsion produite par l'usage de frontières, de seuils, considérant que « l'attention portée aux marges a souvent permis d'économiser la réflexion sur ce qui se passe au centre ». Il prend pour exemple la façon dont l'exclusion est traitée actuellement, qui ne rend compte que la partie émergée de l'iceberg des situations de vulnérabilité vécues par un grand nombre de nos concitoyens.

Ce traitement à la marge d'un problème général est caractéristique de ces politiques sociales. (...) Prendre la question dans sa totalité serait plus coûteux intellectuellement mais aussi pratiquement et politiquement, parce que cela conduirait à mobiliser des politiques de l'emploi plus énergiques que celles qui se réduisent au traitement social du chômage.

- 77 Aussi, préconise-t-il de considérer les marges en référence au centre sans établir de ligne de démarcation avec le souci de délimiter pour compter tout en nous recommandant de garder en mémoire qu'avec la prolifération des marges, c'est la majorité qui risque de devenir déviante.

BIBLIOGRAPHIE

Beck François (en cours). Entre représentativité des échantillons et représentations des usages de drogues. Enjeux, apports et limites des enquêtes en population générale. Thèse en cours.

Beck François, Legleye Stéphane, Peretti-Watel Patrick (2005). Aux abonnés absents : liste rouge et téléphone portable dans les enquêtes en population générale sur les drogues. *Bulletin de Méthodologie Sociologique*, 86 (avril).

Becker Howard S. (1985). *Outsiders. Etudes de sociologie de la déviance*. Paris, A.M. Métailié.

Boltanski Luc, Maldidier Pascale (1977). *La vulgarisation scientifique et son public*. Paris, Rapport du Centre de Sociologie de l'Education et de la Culture, Ehess.

Bon Michel, d'Arc Antoine (1974). *Rapport sur l'homosexualité de l'homme*. Paris, Editions Universitaire (Encyclopédie universitaire).

Bulmer Martin, Bales Kevin, Kish Sklar Kathryn (eds). (1991). *The Social Survey in Historical Perspective, 1880-1940*. Cambridge, Cambridge University Press.

Cassan Francine(coord.) (2002). *Histoire familiale des hommes détenus*. INSEE, collection Synthèses, 59.

Castel Robert (1995a). *Les métamorphoses de la question sociale. Une chronique du salariat*. Paris, Fayard.

Castel Robert (1995b). L'avènement d'un individualisme négatif, propos recueillis par François Ewald. *Magazine Littéraire*, 334 : 18-22.

Cavailhès Jean, Dutey Pierre, Bach-Ignasse Gérard (1984). *Rapport Gai. Enquête sur les modes de vie homosexuels*. Paris, Editions Persona.

Chartier Roger (1989). Le monde comme représentation. *Annales E.S.C.*, 6 : 1513-14.

Chevallier Emmanuelle (2001). *Estimations locales de la prévalence de l'usage d'opiacés et cocaïne en France. Une étude multicentrique à Lens, Marseille, Nice et Toulouse*. Paris, Observatoire français des drogues et des toxicomanies

Coeffic Nicole (1993). L'enquête post-censitaire de 1990. Une mesure de l'exhaustivité du recensement. *Population*, 6, 1655-1680.

Courbage Youssef (1996). Rivalités démographiques en Irlande du Nord. *Population & Société*, 309 : 2-3.

Damon Julien (2002). Les « S.D.F. », de qui parle-t-on ? Une étude à partir des dépêches AFP. *Population-F*, 57(3) : 569-582.

Desrosières Alain (1993). *La politique des grands nombres. Histoire de la raison statistique*. Paris, La Découverte.

Desrosières, Alain et Thévenot, Laurent (1992). *Les catégories socio-professionnelles. La découverte* : Paris.

Fay Robert E., Turner Charles F., Klassen Albert D., Gagnon John H. (1989). Prevalence and Patterns of Same-Gender Sexual Contact Among Men. *Sciences*, 243 (4889) : 338-348.

- Goffman Edwin (1975). *Stigmate, les usages sociaux des handicaps*. Paris, Editions de Minuit.
- Hacking Ian (1986). *Making up people*. In *Reconstructing individualism*. Palo Alto CA, Stanford University Press.
- Jaspard Maryse et l'équipe Enveff (2001). *Nommer et compter les violences envers les femmes : une première enquête nationale*. *Population et Sociétés*, 364 : 1-4.
- Lévy Michel Louis, (1996). *Décrire les minorités ?* *Population & Sociétés*, 309 : 1-2.
- Lévy Michel Louis, (1986). *Nommer pour compter*. *Population & Société*, 202 : 1-4.
- Leyland Alastair, Barnard Marina, McKeganey Neil (1993). *The use of Capture-Recapture Methodology to Estimate and Describe Covert Population: An Application to Female Street-Working Prostitiution in Glasgow*. *Bulletin de Méthodologie Sociologique*, 38 : 52-73.
- Lhomond Brigitte (1997). *Le sens de la mesure. Le nombre d'homosexuel/les dans les enquêtes sur les comportements sexuels et le statut de groupe minoritaire*. *Sociologie et Sociétés*, XXIX(1) : 61-69.
- Marpsat Maryse, Firdion Jean-Marie (2000). *La rue et le foyer. Une recherche sur les sans domicile et les mal-logés dans les années 1990*. Paris, Ined (Travaux et documents, cahier 144).
- Mayer Nonna (2002). *La consistance des opinions*. In Grunberg Gérard, Mayer Nonna, Sniderman Paul M. *La démocratie à l'épreuve. Une nouvelle approche de l'opinion des Français*. Paris, Presse de Sciences Po : 19-49.
- Messiah Antoine, Mouret-Fourme Emmanuelle (1993). *Homosexualité, bisexualité : éléments de socio-biographie sexuelle*. *Population-F*, 48(5) : 1353-1379.
- Mormiche Pierre (1998). *L'enquête HID de l'Insee. Objectifs et schéma organisationnel*. *Courrier des statistiques*, 87-88.
- Pollak Michael, Schiltz Marie-Ange (1991) : *Questions de méthode Six années d'enquête sur les homo- et bisexuels masculins face au sida*. *Bulletin de Méthodologie Sociologique*, 31 : 32-48.
- Pottier Marie-Lys, Robert Philippe, Zauberman Renée (2002). *Victimation et insécurité en Ile-de-France. Les résultats de la première enquête (2001). Rapport final*. Paris, Guyancourt, IAURIF/ CESDIP.
- Ravaud Jean-François, Letourmy Alain, Ville Isabelle (2002). *Les méthodes de délimitation de la population handicapée : l'approche de l'enquête Insee Vie quotidienne et santé*. *Population-F*, 57 (3) : 541-566.
- Richardson Clive (1996). *Leçons à tirer des études de populations animales*. Paris, Séminaire George Pompidou-OFDT.
- Robert Philippe, Pottier Marie-Lys, Zauberman Renée (2003). *Les enquêtes de victimisation et la connaissance de la délinquance. Séminaire de Méthodes d'enquêtes INED-Société française de statistique du 11 avril 2002 (publié en partie dans le Bulletin de Méthodologie Sociologique*, 80 : 5-24.
- Rogers Susan M., Gribble James N., Turner Charles F., Miller Heather G. (1999). *Entretiens autoadministrés sur ordinateur et mesure des comportements sensibles*. *Population-F*, 54(2) : 231-250.
- Rosanvallon, Pierre (1995). *La nouvelle question sociale. Repenser l'état -providence*. Paris, éditions du Seuil.

Schiltz Marie-Ange (1997). Parcours de jeunes homosexuels dans le contexte du VIH : La conquêtes de modes de vie. *Population-F*, 52(6) : 1485-1538.

Schiltz Marie-Ange, Jaspard Maryse, Brown Elizabeth, Firdion Jean-Marie et l'équipe Enveff (2001) : Enquête nationale sur les violences envers les femmes en France (Enveff) : mise en place d'une enquête sur un sujet sensible qui peut mettre en danger les femmes interrogées. in Droesbeke J-J, Lebart L (eds.) *Enquêtes, modèles et applications*. Paris, Dunod, 2001 : 219-230.

Spira Alfred, Bajos Nathalie et le groupe ACSF (1993) : *Les comportements sexuels en France*. Paris, La Documentation française.

ANNEXES

Liste et date de passation des enquêtes retenues pour l'analyse

Rapport sur l'homosexualité de l'homme, enquête auprès des homosexuels masculins, 1972 ;

Rapport gai : enquête sur les modes de vie homosexuels, 1982 ;

Enquête post-censitaire de 1990 ;

Enquête ACSF : Analyse des comportements sexuels en France, 1992 ;

Etude multicentrique sur les attitudes et les comportements des toxicomanes face au risque de contamination par le VIH et les virus de l'hépatite, 1987 et 1990-91 ;

Enquête RMI : sur le devenir des bénéficiaires du RMI , 1997-98 ;

Enquête SDF-Ined auprès des sans-domicile utilisant les services parisiens, 1995 ;

Enquête « Histoire familiale », volet « hommes incarcérés », 1999 ;

Enquête VQS-HID : enquête « Vie quotidienne et santé » liée au recensement de la population de 1999 qui sert de filtrage pour la sélection des individus inclus dans l'enquête HID ;

Enquête HID-ménages : handicaps-incapacités-dépendance auprès des ménages, 1999 ;

Etude multicentrique des estimations locales de la prévalence de l'usage d'opiacés en cocaïne en France, 1998-1999 ;

Enquêtes Presse Gay (EPG): enquêtes par voie de presse auprès des homosexuels masculins répétées de 1985 à 2000 ;

Enquête Enveff : enquête nationale sur la violence envers les femmes en France, 2000 ;

Enquêtes Cesdip-IdF : enquêtes de victimation en Région Ile-de-France, 2001, 2003 ;

Enquête Vespa, VIH : enquête sur les personnes atteintes, 2003.

NOTES

1. Cette première approche ne prend pas en compte l'idée de suivi qui permet de déterminer des flux d'entrée et de sortie des marges.

2. D'abord, à l'occasion d'un numéro spécial, publié en 1993, sur le thème « Sexualité et sciences sociales » Population 48(5), de la publication, en 2000, des résultats d'une recherche sur les sans-domicile et les mal-logés (Travaux et Documents : cahier 144), puis de deux dossiers : un premier, en 1999, consacré aux « Données et enquêtes sensibles » et un second, en 2001, sur les méthodes de collecte et d'analyse pour les enquêtes sur des thèmes sensibles Population 54(2) et 56(6).
3. Un arsenal de textes législatifs encadre cette problématique.
4. Ce qui est vu comme « ordinaire » reflète une norme sociale.
5. Enquêtes en profondeur réalisées après un recensement auprès d'un sous échantillon de la population afin de vérifier la qualité des données recueillies et d'évaluer les biais de collecte.
6. N'oublions pas que le processus de distinction sociale peut jouer dans un sens positif. Il existe des minorités valorisées comme le « happy few », les élites, le cercle des initiés... Sous le coup d'une politique de discrimination positive, une assignation peut se transformer et être revendiquée, ainsi l'explosion des déclarations d'origine amérindienne entre deux recensements aux Etats Unis liée à la mise en place d'une politique de discrimination positive extrêmement favorable d'accès à l'enseignement universitaire.
7. On peut rappeler, à ce propos, l'étrange destin des gauchers corrigés dans les années cinquante et désormais banalisés.
8. De fait, il est désormais de règle que les associations soient associées dès la conception et la mise en place des enquêtes aussi diverses que les Enquêtes Presse Gay, Enveff, HID, Vespa.
9. Comme le souligne très justement F. Beck dans sa thèse en cours, « le terme de « population générale » ne signifie pas « population totale » : une enquête en milieu scolaire sur les usages des substances psychoactives fait partie des enquêtes en population générale pour les âges auxquels le taux de scolarisation reste élevé, c'est-à-dire jusqu'à 17 ans ».
10. Voir, par exemple, Sira et al., 1993.
11. Ce type d'approche a pour objectifs d'établir des prévalences, de fournir des données de cadrage, tout en se donnant la possibilité de décrire les modes de vie des groupes minoritaires.
12. Les chiffres retenus ici portent sur la population âgée de plus de 16 ans.
13. La base de sondage se limite aux les régions participantes et aux seules maisons d'arrêt et centre de détentions qui regroupent les 9/10ème de la population carcérale.
14. Une liste des « centres-jours » est établie, à savoir l'ensemble des prestations distribuées par un centre donné tel jour d'enquête.
15. D'après une enquête de la Direction des Hôpitaux menée en 1996, 93% de l'ensemble des sujets suivis pour le VIH le sont en milieu hospitalier.
16. 2083 individus ont ainsi été sélectionnés parmi les 10734 détenus présents dans les 28 établissements préalablement tirés au sort.
17. De même, les enquêtes téléphoniques qui doivent passer de l'unité « ménage » à l'individu, s'informent sur le nombre de personnes éligibles dans le ménage, les personnes enquêtées étant affectées de pondérations inégales, fonction de ce nombre.
18. En aucune façon, cette approche ne peut se justifier en population ordinaire qui n'a ni style de vie, ni contre culture à défendre, pour lesquelles les techniques probabilistes sont tout à fait adaptées. Ainsi, malgré plus de 300000 réponses, la fiabilité des résultats de la consultation nationale des jeunes de 1994 est fortement controversée. Dans ce cas, la population « cible », à savoir la jeunesse, est susceptible d'être approchée à partir d'une base de sondage correctement établie. De plus, par son coût, cette opération se place en concurrence avec une approche probabiliste qui aurait pu asseoir ses conclusions généralisables à partir d'une démarche inférentielle.
19. Pour définir des groupes à risque susceptibles d'un traitement social préventif ou curatif, les études épidémiologiques font clairement référence à la population éligible.
20. Le bilan des enquêtes sur les limitations d'activité menées par la Statistique du Canada dans le cadre des recensements de 1986 et 1991 a montré les déficiences d'un filtrage trop sommaire.

Les personnes souffrant de limitation d'activité étaient sélectionnées sur la base de deux questions générales. Un individu sur dix ainsi sélectionné s'est avéré être un « faux positif ». Plus grave, l'enquête de contrôle auprès de personnes non filtrées a révélé que les personnes qui avaient échappé au filtrage de façon injustifiée représentaient plus du tiers des personnes recherchées.

21. Cette technique est issue de l'étude de la taille de populations animales sauvages. L'idée du modèle est fondée sur le tirage d'échantillons d'animaux que l'on tatoue et relâche dans la nature avant d'effectuer d'autres prélèvements. Le pourcentage d'animaux tatoués « recapturés » permet d'estimer la taille de la population.

RÉSUMÉS

Après avoir souligné les divergences de point de vue entre la statistique et la sociologie sur la manière d'appréhender les « populations difficiles à atteindre », cet article se propose d'explorer différentes façons de définir l'écart par rapport à une norme, la périphérie par rapport à un centre et de montrer comment, selon le choix du critère d'inclusion adopté, la taille et le sens du groupe statistique constitué sont extrêmement variables.

Making and Unmaking Groups, Numerical Information on "Difficult-Access Populations":

After having underlined the divergent points of view between statistics and sociology on how to approach "difficult-access populations", this article proposes to explore the different manners of defining a distance from a norm, the periphery in relation to a center, and to show how, according to the choice of criteria of inclusion adopted, the size and meaning of the statistical group thus constituted are extremely variable.

INDEX

Mots-clés : Enquêtes, Populations difficiles à atteindre, Sensitivité méthodologique

Keywords : Hidden Populations, Methodological Sensitivity, Surveys

AUTEUR

MARIE-ANGE SCHILTZ

CAMS-CERMES, schiltz@ehess.fr